

# WEATHER AND PROBABILITY OF OUTBREAKS OF THE PALE WESTERN CUTWORM IN MONTANA AND NEAR-BY STATES<sup>1</sup>

By WILLIAM C. COOK<sup>2</sup>

[Assistant Entomologist, Montana Agricultural Experiment Station, Bozeman, Mont.]

It has been shown by Seamans (1) and by the writer (2) that variations in the population of the pale western cutworm (*Porosagrotis orthogonia* Morr.) may be forecast with a high degree of reliability from the rainfall in the spring of the preceding year. Seamans showed that if there were less than 10 days in May and June upon which 0.25 inch of rainfall was recorded, *orthogonia* would increase in the following year, while if there were more than 10 "wet" days, *orthogonia* would decrease. I have shown, similarly, that if the total rainfall in May, June, and July was less than 4 inches, the population would increase, while if this total was more than 5 inches, the population would decrease.

These two methods of predicting changes in abundance may easily be harmonized. It is well known that there is a definite relation between total rainfall and the frequency of various amounts in single showers. Cole (3) has shown that in the Judith Basin region of Montana less than 25 per cent of the total rainfall from April 1 to September 1 falls in showers of less than 0.20 inch. This relation should hold approximately for May, June, and July. Since Seamans has used only the rainfall for May and June, in which the heaviest rain falls, 3 inches may be taken as the critical total for these two months. Assuming 25 per cent, or 0.75 inch, to fall in light showers, this would leave 2.25 inches—sufficient for eight or nine showers of 0.25 inch each. It seems probable that a careful analysis of long records in this region would bring the two forecasting methods into still closer correspondence.

Statistical studies of over 60 weather records for an outbreak of *P. orthogonia* in Montana from 1918 to 1922 have shown biserial correlations between rainfall and presence of damage as follows:

Period	Correlation coefficient (r)
May of preceding year.....	-0.471 ± 0.105
June of preceding year.....	-.572 ± .092
July of preceding year.....	-.675 ± .053
May and June, total.....	-.730 ± .050
May, June, and July, total.....	-.766 ± .045

A correlation of -0.766 indicates a relation, as measured by  $\sqrt{1-r^2}$ , of 0.643, showing well over half the possible causes accounted for. This relation to rainfall may be regarded as vital, and a calculation of the probability of outbreaks upon this basis should give a maximum value as other neglected factors would probably operate to reduce rather than increase the number of outbreaks.

In the paper previously referred to (2), I have shown that the critical factor is really soil moisture, which is a function of temperature as well as of rainfall. Within the region covered by this study, however, temperatures are very uniform at this time of the year, and would cause very small changes in the critical rainfall of 4 inches.

The problem is thus resolved into a study of the probability that the total rainfall from May 1 to August 1 will be less than 4 inches. This probability measures the adaptability of the climate for continued habitation by *P. orthogonia*, but will not measure the probability of a serious outbreak. Careful field and laboratory studies have shown that even under the most favorable conditions, this species is rarely able to more than treble its population from one year to another. Field studies have shown that the normal population, in cutworm areas, when there is no outbreak, is less than one larva per square yard, while it requires at least six per square yard to produce severe damage. This would indicate that at least two favorable years are necessary to produce an outbreak, which is substantiated by the fact that every severe outbreak studied has been preceded by at least two favorable years.

Thus, as a second necessary factor, the probability of two successive dry years is needed. For these analyses long continuous rainfall records are necessary. Within the Montana area occupied by *P. orthogonia* (4), six such records have been selected (Helena, Havre, Crow Agency, Miles City, Glendive, and Poplar). A seventh long record from Bozeman, just outside the region occupied by *orthogonia*, has been added for comparison. The severe outbreak of *orthogonia* in Montana was accompanied by a widespread but less severe outbreak in western North Dakota, indicating a decrease in severity with increase in normal rainfall. Williston and Bismarck, N. Dak., and Moorhead, Minn., have been added to show the influence of this factor upon the probability of outbreaks. Outbreaks of *orthogonia* have occurred in northeastern Colorado; and Cheyenne, Wyo., and Fort Collins, Colo., represent this area. Unfortunately no long records exist in the small areas covered by very recent outbreaks in northeastern New Mexico and extreme western Oklahoma.

## THE PROBABILITY OF 4 INCHES RAINFALL OR LESS DURING MAY, JUNE, AND JULY AT THE SELECTED STATIONS

As is generally known, rainfall values in arid regions form a skew distribution, in which the mode is usually less than the mean. This is the case in this study, and it is perfectly useless to apply a normal curve. A simple counting of the values below 4 inches would be more accurate. It would be possible to fit a form of Pearson's generalized frequency curve to each of them, but the process is very laborious. Tolley (5) has published an alternative family of frequency curves, especially adapted for this type of work, and has tabulated the probability of deviations in terms of the standard deviation for selected values of skewness. His measure of skewness, K, is the ratio  $\mu_3/\sigma^3$ , so that all necessary constants are derived from the first three moments. His curves have been fitted to the rainfall distributions for the selected stations. These distributions, together with the various constants, are given in Table 1, and the histograms are plotted in Figure 1.

<sup>1</sup> Contribution from the entomology department, Montana Agricultural Experiment Station.

<sup>2</sup> The writer wishes to acknowledge the advice and assistance of Prof. W. D. Tallman, head of the mathematics department, Montana State College, in this and other similar studies.

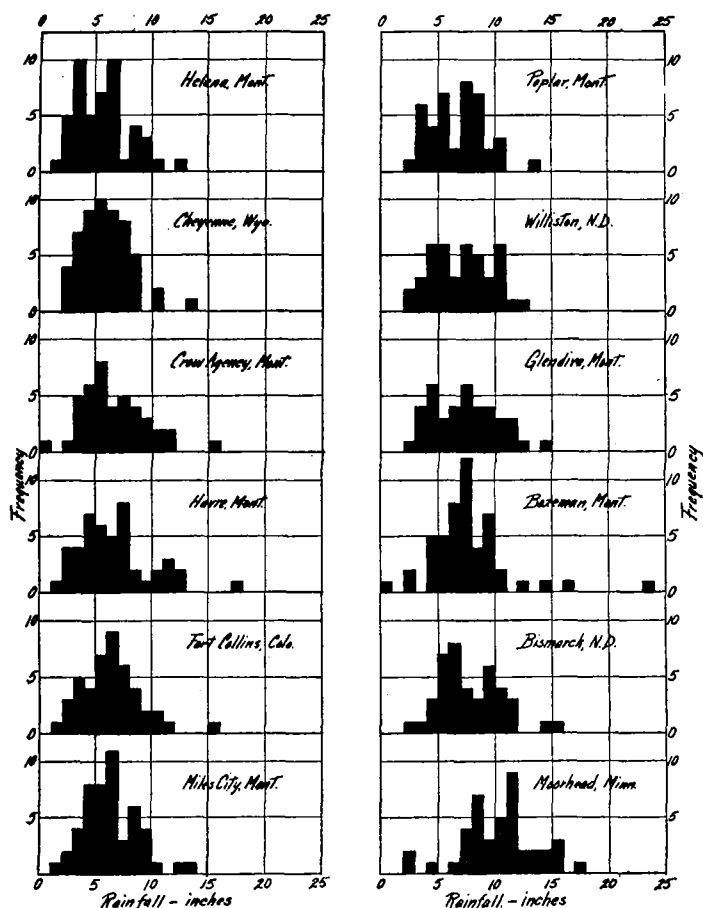


FIG. 1.—Frequency of various amounts of rainfall at the 12 selected stations. (See Table 1)

TABLE 1.—Frequency distributions and statistical constants for May-July rainfall at 12 selected stations

Total rainfall, inches	Helena, Mont.	Cheyenne, Wyo.	Crow Agency, Mont.	Havre, Mont.	Fort Collins, Colo.	Miles City, Mont.
0.00-0.99			1			
1.00-1.99	1			1	1	1
2.00-2.99	5	4	1	4	3	2
3.00-3.99	10	7	5	4	5	4
4.00-4.99	5	9	6	7	4	8
5.00-5.99	7	10	8	6	7	8
6.00-6.99	10	9	4	5	9	11
7.00-7.99	1	8	5	8	6	3
8.00-8.99	4	5	4	2	4	6
9.00-9.99	3		3	1	2	4
10.00-10.99	1	2	2	2	2	1
11.00-11.99			2	3	1	
12.00-12.99	1			2		1
13.00-13.99		1				1
14.00-14.99						
15.00-15.99			1		1	
16.00-16.99						
17.00-17.99				1		
18.00-18.99						
19.00-19.99						
20.00-20.99						
21.00-21.99						
22.00-22.99						
23.00-23.99						
Record, year's	48	55	42	45	45	50
Mean (M)	5.563	5.918	6.595	6.433	6.400	6.400
Standard deviation ( $\sigma$ )	2.332	2.169	2.912	2.825	2.644	2.418
Coefficient of variation	41.9	36.6	44.2	45.3	41.4	37.8
Third moment ( $\mu_3$ )	+6.778	+8.410	+18.362	+11.279	+16.516	+8.760
K	+0.535	+0.824	+0.744	+0.500	+0.893	+0.630
Probability of 4 inches or less (per cent)	26.78	19.39	19.37	19.14	18.89	16.08
Average interval, years	3.74	5.15	5.15	5.23	5.29	6.21

<sup>1</sup> These values were omitted in making the calculations.

TABLE 1.—Frequency distributions and statistical constants for May-July rainfall at 12 selected stations—Continued

Total rainfall, inches	Poplar, Mont.	Williston, N. Dak.	Glendive, Mont.	Bozeman, Mont.	Bismarck, N. Dak.	Moorhead, Minn.
0.00-0.99				1		
1.00-1.99						
2.00-2.99	1	2	1	2	1	2
3.00-3.99	6	3	4		3	
4.00-4.99	4	6	6	5	1	1
5.00-5.99	7	6	3	5	7	
6.00-6.99	2	3	4	8	8	1
7.00-7.99	8	6	6	12	4	4
8.00-8.99	7	5	4	4	3	7
9.00-9.99	2	3	4	7	6	3
10.00-10.99	3	6	3	2	4	5
11.00-11.99		1	3		3	9
12.00-12.99		1	1	1		2
13.00-13.99	1					2
14.00-14.99			1	1	1	2
15.00-15.99					1	3
16.00-16.99				1		
17.00-17.99						1
18.00-18.99						
19.00-19.99						
20.00-20.99						
21.00-21.99						
22.00-22.99						
23.00-23.99				1		
Record, year's	41	42	40	50	42	42
Mean (M)	6.744	7.095	7.375	7.398	7.833	10.357
Standard deviation ( $\sigma$ )	2.410	2.583	2.846	2.767	2.759	3.213
Coefficient of variation	35.7	36.5	38.5	37.3	35.2	31.1
Third moment ( $\mu_3$ )	+5.138	+1.733	+8.221	+13.271	+13.026	-8.532
K	+0.367	+0.100	+0.357	+0.633	+0.630	-0.257
Probability of 4 inches or less (per cent)	12.33	11.37	11.23	9.94	6.75	3.01
Average interval, years	8.11	8.80	8.90	10.10	14.81	3.32

<sup>1</sup> These values were omitted in making the calculations.

From a climatological standpoint, the table calls for a few comments. The coefficient of variation ( $100\sigma/M$ ) shows a fairly regular decrease from west to east, as may be seen by comparing Havre (45.3), Miles City (37.8), Williston (36.5), Bismarck (35.2), and Moorhead (31.1). This is roughly parallel to the increase in total rainfall.

The skewness of the distribution is positive in every case except Moorhead, and negative skewness is obvious in the histogram for that station. Williston has a freakish distribution, with little skewness, showing nearly equal frequencies in the central classes. Poplar and Glendive, the nearest Montana stations, also show this property in a less pronounced form. It is difficult to predict just what sort of a curve would be obtained from longer records. These stations are directly in the path of the Alberta cyclones in summer, and this may possibly be a factor in producing the irregular curves. Elsewhere in the table, the values of  $K$  for the Northern States fall near +0.600, and in Colorado near +0.800.

The values of the probability ( $P$ ) of 4 inches rainfall or less, obtained by interpolation from Tolley's table (5) p. 640-641) show a wide variation. This probability gives a clue to the conditions which determine the distribution of this species. *P. orthogonia* is known to maintain a small population near Miles City ( $P=16.08$ ) and does not maintain itself near Bozeman ( $P=9.94$ ). This would indicate that a limiting value would be slightly above  $P=10$ . The species has migrated, in favorable years, into the region near Bismarck ( $P=6.75$ ), but has never been reported near Moorhead ( $P=3.01$ ). This indicates that outbreaks caused by migrating moths would probably occur within lines indicating values of  $P$  between 5 and 10.

Another way of conveying the same information is by means of the reciprocal of  $P$ , or the average interval

between favorable years. The average intervals are given in the lowest line of Table 1. Translating the above paragraph into terms of average interval, it may be stated that *P. orthogonia* can maintain itself at all times only in regions where at least 1 year in 10 is favorable for an increase. It can produce sporadic outbreaks, by migration, in regions where more than 1 year in 20 is favorable. This latter interval marks the extreme limit of the economic distribution. Only isolated specimens will be found elsewhere.

#### THE PROBABILITY OF TWO SUCCESSIVE FAVORABLE YEARS

According to the laws of chance, if the chance of a single favorable year is  $P$ , then that of two successive favorable years is  $P^2$ , assuming a random (normal) distribution of rainfall values. A large amount of work has appeared recently which attempts to discern cycles in the values of weather elements. It is not within the scope of this paper to discuss any possible cycles, but a mere inspection of any rainfall record from the Great Plains area will show a decided tendency for like years to succeed each other. There is a very simple method of evaluating this tendency, while avoiding any implication of cyclic action. If there is no tendency for like years to succeed each other, the differences between successive years, taken *without regard to sign*, should form a normal distribution. If such a tendency be present, there should be an accumulation of small differences, producing a curve of positive skewness. From this difference curve, following the method of Tolley, the probability of any desired degree of similarity may be calculated. Since I have shown (2) that the critical range is 4 to 5 inches, it is sufficient to determine the probability of a deviation of 1 inch or less. Distributions of the interannual differences for the selected stations, with their constants, and the probability of a deviation of 1 inch or less are given in Table 2.

TABLE 2.—Frequency distribution and statistical constants for interannual differences in May-July rainfall

Difference in inches	Helena, Mont.	Cheyenne, Wyo.	Crow Agency, Mont.	Havre, Mont.	Fort Collins, Colo.	Miles City, Mont.
0.00-0.49	8	3	5	7	2	5
0.50-0.99	6	7	4	6	2	8
1.00-1.49	4	10	3	3	4	5
1.50-1.99	6	7	4	5	10	3
2.00-2.49	3	5	5	1	3	4
2.50-2.99	5	12	4	2	1	4
3.00-3.49	2	5	1	3	4	6
3.50-3.99	2	1	3	1	2	6
4.00-4.49	2	1	3	4	3	2
4.50-4.99	6	1	1	1	1	2
5.00-5.49	1	2	2	4	3	1
5.50-5.99	1	1	1	1	1	1
6.00-6.49	1	1	1	1	1	1
6.50-6.99	1	1	1	1	1	1
7.00-7.49	1	1	1	1	1	1
7.50-7.99	1	1	1	1	1	1
8.00-8.49	1	1	1	1	1	1
8.50-8.99	1	1	1	1	1	1
9.00-9.49	1	1	1	1	1	1
9.50-9.99	1	1	1	1	1	1
10.00-10.49	1	1	1	1	1	1
10.50-10.99	1	1	1	1	1	1
11.00-11.49	1	1	1	1	1	1
11.50-11.99	1	1	1	1	1	1
12.00-12.49	1	1	1	1	1	1
12.50-12.99	1	1	1	1	1	1
Year's	40	54	38	43	41	49
Mean difference	2.388	2.242	3.289	3.128	3.329	2.587
Standard deviation ( $\sigma$ )	1.792	1.498	2.652	2.670	2.568	1.861
Coefficient of variation	75.0	66.7	80.6	85.4	77.2	71.9
Third moment ( $\mu_3$ )	+3.511	+5.102	+23.555	+16.798	+27.401	+5.588
K	+0.610	+1.519	+1.264	+0.883	+1.618	+0.089
Probability of a difference of 1 inch or less	23.62	22.56	23.20	22.83	17.69	20.49

TABLE 2.—Frequency distribution and statistical constants for interannual differences in May-July rainfall—Continued

Difference in inches	Poplar, Mont.	Williston, N. Dak.	Glendive, Mont.	Bozeman, Mont.	Bismarck, N. Dak.	Moorhead, Minn.
0.00-0.49	2	6	2	4	3	2
0.50-0.99	3	1	4	7	7	2
1.00-1.49	3	5	2	2	4	4
1.50-1.99	3	2	6	4	1	5
2.00-2.49	4	3	3	7	2	1
2.50-2.99	3	6	2	6	3	3
3.00-3.49	3	1	3	1	1	4
3.50-3.99	1	5	4	1	7	1
4.00-4.49	2	4	2	2	5	4
4.50-4.99	4	3	2	1	3	3
5.00-5.49	3	1	1	3	1	2
5.50-5.99	1	1	1	2	1	1
6.00-6.49	1	1	1	1	1	2
6.50-6.99	1	2	1	1	1	2
7.00-7.49	2	1	1	1	1	1
7.50-7.99	1	1	1	1	1	1
8.00-8.49	1	1	1	1	1	3
8.50-8.99	1	1	1	1	1	1
9.00-9.49	1	1	1	1	1	1
9.50-9.99	1	1	1	1	1	1
10.00-10.49	1	1	1	1	1	1
10.50-10.99	1	1	1	1	1	1
11.00-11.49	1	1	1	1	1	1
11.50-11.99	1	1	1	1	1	1
12.00-12.49	1	1	1	1	1	1
12.50-12.99	1	1	1	1	1	1
Year's	35	41	38	44	41	41
Mean difference	3.243	2.970	3.421	3.023	2.988	3.988
Standard deviation ( $\sigma$ )	1.954	1.973	2.511	2.609	2.017	2.608
Coefficient of variation	60.3	66.5	73.4	86.3	72.6	65.4
Third moment ( $\mu_3$ )	+3.115	+4.033	+1.916	+26.536	+4.685	+10.639
K	+0.417	+0.525	+0.968	+1.496	+0.565	+0.600
Probability of a difference of 1 inch or less	12.05	15.87	16.36	24.74	16.21	11.84

With the exception of Bismarck and Miles City, the stations show a markedly more skew distribution of differences than that for the original records. In three cases extrapolation was necessary, as Tolley's table extends only to  $K = +1.4$ . In the cases of Helena and Havre, the fit of Tolley's curve was forced, as the highest frequencies are concentrated in the first two classes, showing almost a hyperbolic relationship. As the records are comparatively short, there seemed to be no object in attempting to get a better fit with some other curve. These variations, however, did have a decided effect upon the relation of actual to calculated probability. Bozeman has a very skewed curve and a wide variability.

The values of  $P$  at the bottom of the table indicate only the probability of two like years, regardless of the absolute value. They must be considered in connection with the values of  $P$  for the probability of a single year with less than 4 inches rainfall. Both values of  $P$ , with the compound probability (calculated from  $P_1^2$  and also from  $P_1 \times P_2$ ), the average interval between outbreaks, and the actual and computed numbers of pairs of dry years in each record, are given in Table 3.

TABLE 3.—The probability of cutworm outbreaks

Column	1	2	3	4	5	6	7	8
Station	Length of record, years	Chance of less than 4 inches rainfall ( $P_1$ )	Chance of less than 1 inch deviation ( $P_2$ )	Chance of 2 successive years, calculated from— ( $P_1$ ) <sup>2</sup>	( $P_1$ ) $\times$ ( $P_2$ )	Average interval, years	Pairs of years with less than 4 inches rainfall each— Expected	Found
Helena, Mont.	48	26.78	23.62	7.2	6.3	16	3.02	6
Cheyenne, Wyo.	55	19.39	22.56	3.8	4.4	23	2.42	3
Crow Agency, Mont.	42	19.37	23.20	3.7	4.5	22	1.89	1
Havre, Mont.	45	19.14	22.53	3.7	4.4	23	1.98	2
Fort Collins, Colo.	45	18.89	17.69	3.6	3.3	30	1.49	1
Miles City, Mont.	50	16.08	20.49	2.6	3.3	30	1.66	0
Poplar, Mont.	41	12.33	12.05	1.5	1.5	67	0.61	1
Williston, N. Dak.	42	11.37	15.87	1.3	1.8	55	0.76	2
Glendive, Mont.	40	11.23	16.36	1.3	1.8	55	0.72	1
Bozeman, Mont.	49	9.94	24.74	0.9	2.5	40	1.22	0
Bismarck, N. Dak.	42	6.75	16.21	0.5	1.1	91	0.48	0
Moorhead, Minn.	42	3.01	11.84	0.1	0.4	250	0.17	0

As all of these calculations are based upon relatively short records, it did not seem wise to give a false impression of accuracy by carrying the final probability values to many decimals, and the average intervals are given to the nearest whole year.

Column 6 of this table shows that severe outbreaks of *P. orthogonia* may be expected about once in 20 years in central Montana and Wyoming, about once in 30 years in eastern Montana and Colorado, and about once in 50 years in North Dakota. The interval of 250 years at Moorhead places this station definitely outside the economic distribution of this insect. It is, of course, possible that one extremely favorable year following one with a rainfall between 4 and 5 inches may cause a local outbreak of moderate severity, and such local outbreaks may be expected somewhat more frequently. There are indications, not climatic in character, that a small district in southern Alberta and extreme north-central Montana may suffer more frequent outbreaks than any other region studied, but no long rainfall records are available for a climatic study.

The comparison of actual and expected pairs of dry years shows a fairly close fit of theory to observation. The wide deviation in the case of Helena is explained above, while in the case of Crow Agency there are several gaps in the record which might have included one or more dry years.

#### CONCLUSIONS

This study has shown that there is a close connection between the chances of a single favorable year and the ability of *Porosagrotis orthogonia* to maintain itself at a place. At least 1 year in 10 must be favorable to increase

or the insect will disappear. Outbreaks brought on by migrating moths are confined to regions where at least 1 year in 20 is favorable.

There is a decided tendency for like years to follow in succession, and this tendency was evaluated, without any implication of periodicity, by forming a distribution curve of interannual differences, neglecting signs. In all cases a decidedly skewed distribution was obtained from which the probability of a deviation of 1 inch or less between successive years was calculated.

Severe attacks of the pale western cutworm may be expected only at long intervals in most parts of its range. In the dry mountain foothills near Helena such outbreaks will be about 16 years apart, on the central plains of Montana and in Wyoming it may occur once in 20 to 25 years, and in eastern Montana and Colorado the average interval between outbreaks is about 30 years. In other parts of its range in the United States it will probably be more rarely injurious. Light outbreaks may be expected somewhat more frequently.

#### LITERATURE CITED

1. Seamans, H. L. 1923. Forecasting outbreaks of the pale western cutworm in Alberta. *Can. ent.* 55: 51-53.
2. Cook, W. C. 1926. Some weather relations of the pale western cutworm (*Porosagrotis orthogonia* Morr.). A preliminary study. *Ecology* 7: 37-47.
3. Cole, J. S. 1922. The daily quantities in which summer precipitation is received. *Mon. weather rev.* 50: 572-575.
4. Cook, W. C. 1924. The distribution of the pale western cutworm, *Porosagrotis orthogonia* Morr.: A study in physical ecology. *Ecology* 5: 60-69.
5. Tolley, H. R. 1916. Frequency curves of climatic phenomena. *Mon. weather rev.* 44: 634-642.

### ON THE MEASURE OF CORRELATION: A REJOINDER

By GILBERT T. WALKER

The note upon my article on the above subject by Mr. Edgar W. Woolard in the MONTHLY WEATHER REVIEW of October last, raises several points of interest, and I hope I may be allowed to make some comments.

1. The fundamental issue can best be seen in a numerical example. The longest series of forecasts known to me is that based on a formula of 1908 for monsoon forecasting in India. The departures of monsoon rainfall given on the 1st of June of the years 1909-1927 by the regression equation, i. e., the amounts determined by external factors, were:

1909-----	+1.4	1919-----	-0.7
1910-----	+1.3	1920-----	-2.0
1911-----	+0.7	1921-----	+2.6
1912-----	+2.1	1922-----	-0.2
1913-----	-0.6	1923-----	+1.5
1914-----	+0.7	1924-----	+0.5
1915-----	-1.0	1925-----	+0.8
1916-----	+4.5	1926-----	+1.1
1917-----	+2.9	1927-----	+1.4
1918-----	+0.2		

The actual departures during these years were:

1909-----	+1.9	1919-----	+3.2
1910-----	+2.0	1920-----	-4.3
1911-----	-3.2	1921-----	+1.4
1912-----	-1.1	1922-----	+2.0
1913-----	-1.7	1923-----	+1.5
1914-----	+3.4	1924-----	+3.1
1915-----	-3.0	1925-----	-1.6
1916-----	+5.0	1926-----	+3.6
1917-----	+7.1	1927-----	-1.6
1918-----	-6.5		

2. The S. D. of the first series is 1.73, and of the second is 3.44, the ratio of the first S. D. to the second being 0.50; while we find the correlation coefficient between the two series is 0.56. Now  $r^2$  is 0.31, and I contend that in the long run the ratio of the first series to the second will be  $r$ , not  $r^2$ . This fraction must, in the long run, be the same in whatever reasonable way the two series are compared, for they obey the same distribution law. When it is said that I take the S. D. as the measure of variation, while Krichewsky takes the square of the S. D., the statement is, in effect, precisely equivalent to saying that I compare the actual series while Krichewsky compares their squares. I hold that if we compare 1, 2, 3, 4, 5, . . . with 2, 4, 6, 8, 10, . . . the first series is half the second and that we can not justify calling it a quarter (or an eighth) by saying that the square (or the cube) of the terms is considered.

3. I would like now to comment on some of the arguments used. When I state that a fraction  $r\sigma_0$  of  $\sigma_0$  is due to variations in  $x_1$ , I do not attribute "the remainder  $(1-r)\sigma_0$  to variations in  $x_2$ , . . ." (p. 460). For if  $x_0 = p + q$ , where  $p$  and  $q$  are independent, it is easily seen that the S. D.'s  $\sigma_p$ ,  $\sigma_q$ , of  $p$ ,  $q$ , satisfy the equation  $\sigma_p^2 + \sigma_q^2 = \sigma_0^2$ , not  $\sigma_p + \sigma_q = \sigma_0$ ; so the S. D. of the remainder here must be  $\sigma_0(1-r^2)^{1/2}$ .

On p. 461 there is a determination of the average value of the term  $r\sigma_0 \frac{x_1}{\sigma_1}$ , which is accepted as the contribution from  $x_1$ . Now I gave in paragraph 4 on p. 460